COMPARING THREE INPUT DEVICES FOR SKETCHING ASSIGNMENTS IN E-EXAMS IN COMPUTER SCIENCE

Suhas Govind Joshi and Live Årmot Brastad
Department of Informatics, Faculty of Mathematics and Natural Sciences, University of Oslo, Norway

ABSTRACT

E-exams used in higher education can accommodate a range of assignment types that have traditionally been conducted with pen-and-paper. However, one issue that remains a challenge is proper support of sketching in larger, generic e-exams systems. Different types of sketching assignment demand input devices suited to the specific type of problem-solving. Adopting an e-exam system may involve compromises in terms of restricted opportunities to use sketching assignments as part of the e-exam that affect the students' performance or their experience of the e-exam. This paper presents a comparison of 77 students' performance and preferences with three different input devices (a mouse, a graphical tablet, and a touchpad) when solving sketching assignments from three different computer science courses. The study focused on capturing the students' perspective, and the findings include identified benefits and challenges associated with using the three different input devices during sketching.

KEYWORDS

Sketching, e-Exam, Graphical Tablet, Touchpad

1. INTRODUCTION

The increasing presence of e-exams in higher education has brought forward new opportunities to organize exams in ways that were not possible with traditional pen-and-paper exams. Sketching has always been a part of how students are tested in different courses within higher education. Visual communication is an important part of how students attain and convey their knowledge in many learning situations, including e-exam situations. However, shifting sketching assignments that have traditionally been conducted with pen-and-paper over to larger, generic e-exam systems remains a challenge. While e-exams have opened up new possibilities that may replace old sketching assignments with other pedagogical alternatives, technical restrictions or inefficiencies should not prevent sketching assignments from appearing in exam situations.

This paper reports from a study on how students experience sketching assignments as part of e-exams. We have conducted a comparison of performance and preference when students use three different input devices, namely a mouse, a graphical tablet, and a touchpad, to solve representative sketching assignments from computer science courses using an e-exam software. 77 students participated in testing these three input devices in three different undergraduate courses within computer science that build of different traditions and represent different types of sketching assignments.

The study aimed to investigate differences in performance and preference within each course rather than make an overarching cross-course comparison where the sketching assignments were normalized in complexity and degree of freedom. Our underlying assumption was that the different courses would yield different performances with the three input devices due to the variation in complexity and degree of freedom between the representative sketching assignments from each course. We argue that normalizing these assignments would compromise on important details in each course and no longer be representative of how the students experience them during real e-exam situations. As such, the study was a bottom-up comparison where the performance time and preferences with each input device were mainly considered within each course. However, we also studied how the rank and order within each course compared across the three courses to answer the overarching two-folded research question. The two objectives that the research question aimed to answer were: (1) does one input device perform unanimously better in completion time and preference for sketching assignments across the three courses, and (2) what are the main benefits and challenges associated

with the three input devices when used in sketching assignments? Throughout the study, our focus remained on answering these questions from the students' perspective exclusively rather than attempt to include other important end-users of the e-exams systems such as the teachers (Osang, 2012).

To answer our overarching research question, we collected both quantitative and qualitative data to analyze performance and preference. The quantitative data was gathered by measuring the completion time in a randomized experiment and analyzed using a two-way factorial ANOVA. The analysis was complemented with qualitative data from post-test interviews with all participants. Based on the analysis, we present an overview of main benefits and challenges associated with each of the three input devices. The discussion aligns our findings with the literature on students' perspective on e-exams and e-exams including sketching assignments.

2. RELATED WORK

There is a multitude of studies concerning the use of e-examination and e-assessment. The shift to e-exams has been motivated by a range of underlying practical and economic concerns, e.g., streamlining of academic assessment to avoid paper-based work (Ibrahim et al., 2015, p. 48), reducing time spent on the assessment (Kuikka et al., 2014, p. 4; Wibowo et al., 2016), preventing cheating, environmental concerns (Wibowo et al., 2016, p. 18), and for security and authentication reasons (Abdel Karim and Shukur, 2016). Even digital learning technologies with student-focused overarching goals or innovative approaches to exploration with new interaction techniques and tools have yielded varying results (Drake et al., 2015, pp. 125–126).

Most studies on e-exams emphasize the students' view, and many such studies advocate the importance of localizing to support local cultures in teaching specific topics (Lazem, 2016). Examples of strategies for localization within the context of e-exams include building on existing practice in teaching (Oviatt et al., 2006, p. 1), incorporating learning modalities from early childhood to stimulate familiarity (Rante et al., 2018, p. 2), or bringing in familiar technologies to increase the overall learning potential (Romero et al., 2009, p. 445; Wibowo et al., 2016, p. 8). While our study is limited to the perspective of the students, the issue of considering all end-users have been emphasized by multiple studies. For instance, adapting the design of the learning technologies to include concerns at individual, social, and cultural level is an important part of addressing higher-order issues such as inclusion and accessibility (Drake et al., 2015, p. 133). Other examples of relevant considerations include *"aspects of flexibility, clarity, security, and convenience"* (Wibowo et al., 2016, p. 26) and the usability concerns of the teachers during the introduction of a digital system (Kuikka et al., 2014, pp. 5–6).

Prior work of particular interest to our research are those studies concerned with either how students perceive the digital input devices as part the digitalization of exams, or that examine the use of drawing or sketching as the part of the e-examination or a larger learning environment.

2.1 Digital Input Devices and Sketching

Learning systems that incorporate various digital input devices have also been studied in the past. Several studies have examined the role of alternative input devices in learning situations to enhance learning, e.g., studies applying tablets or touch- or pen-based interfaces, (Barreto et al., 2016; Koile and Rubin, 2016; Lara-Garduno et al., 2016; Polsley et al., 2016; Wang et al., 2016; Williford et al., 2016). These studies have reported on various aspects of sketching such as benefits of shifting from paper-and-pencil tests to digital alternatives (Barreto et al., 2016, p. 143; Lara-Garduno et al., 2016, p. 166), automation of interpretation of freehand sketching (Williford et al., 2016, p. 117), performance and behavior in problem solving (Polsley et al., 2016, p. 110), and how digitalizing teaching with tablet-based technology may affect conceptual understanding (Koile and Rubin, 2016, p. 72). Studies have also been done where different input devices have been compared (Oviatt et al., 2006; Thompson et al., 2016).

Sketching as a part of learning and teaching is not something new, and the idea of digitalizing sketching was discussed already two decades ago (Landay and Myers, 2001). For the longest time, sketching has been a natural part of examinations, which were mainly conducted on pen-and-paper that did not introduce any technical challenges. Studies have discussed the importance of sketching within different traditions adopted or embraced by computer science courses, e.g., engineering (Williford et al., 2016; Yang and Cham, 2006),

mechanics (Polsley et al., 2016), and mathematics (Oviatt et al., 2006). One of the arguments of replacing paper exams with e-exams is the support of multimedia content (Wibowo et al., 2016, p. 6). While not being specifically limited to sketching-based exams, studies of performance in digital assessments versus pen-and-paper assessment have been mixed, and in certain exam formats, yielded small differences that were not practically significant (Walker and Handley, 2016, p. 2). In studies of preference rather than performance, similar equally divided results have been reported (Wibowo et al., 2016, p. 16). Comparisons between different input devices have also suggested that the performance of stylus and touch did not vary to any statistically significant degree (Thompson et al., 2016, p. 67). Nevertheless, the challenge of retaining the benefits of paper while digitalizing the interaction has motivated many relevant studies on the use of novel digital interfaces, e.g., with the use of graphical tablets or digital stylus (Lee et al., 2017; Oviatt et al., 2006; Rante et al., 2018; Thompson et al., 2016; Williford et al., 2016). Other examples of state-of-the-art concepts building on experiences with paper include thin-film display technologies (Klamka and Dachselt, 2017) or augmented paper interfaces (Tsandilas et al., 2015).

2.2 Defining Sketching

In the midst of digitalization, sketching remains a tacit skill that enables visual communication during exams. There are many definitions of sketching, and it carries different meanings in different contexts (Lewis et al., 2019, p. 2). The literature also contains ambiguities (Lane 2017, p. 5). Sketching serves multiple roles in a learning situation, and different problems involve sketching in different capacities. Exam systems with tools that support any form of drawing are usually well received, e.g., as seen in (Wibowo et al., 2016, p. 22). However, the understanding of what such tools need to do to support sketching varies greatly from one context to another. The call for automatic correction of multimedia content in e-exam systems has been reported (Kuikka et al., 2014, p. 10), but such content usually refer to videos and pictures. From our perspective, tools offering simplistic self-composed graphic from a predefined array of elements or supporting drawing of a simple line would only constitute as *"drawing with a pen"* (Do, 2002, p. 151), rather than actually support sketching. For instance, we consider freedom and flexibility to remain an important characteristic of sketching (Landay and Myers, 2001, p. 56). The understanding of sketching in our study aligns with how Lane defines sketching as a *"fundamental skill that supports higher-order cognitive processes such as creative problem-solving and innovative critical-thinking"* (2017, p. 2) and that can also serve as a thinking tool for both spatial and visual reasoning. During an exam situation, sketching serves as a part of the reflection of the cognitive activities of the student and is valuable for communicating ideas (Yang and Cham, 2006, p. 476). The sketching assignments included in our study suit the definition of *"study sketches"* and the definition of freehand sketching as a *"reflexive process in which externalizations are communicated"* (Lane, 2017, p. 7) aligns well with how the students answer them.

3. RESEARCH METHOD

This study included 77 students who participated in solving sketching assignments from three different courses where performance times were measured and then followed up by a post-test interview. As our overarching goal was to conduct a bottom-up study where each course was considered individually, we were not interested in the absolute values of the performance, but rather the relative differences between the three input devices observed in the three courses included. The comparison between the three courses was only made to assess the relative rank and order in performance time and preference rather than attempt to make independent comparisons of absolute values.

3.1 Courses and Sketching Assignments

We included representative sketching tasks from three different entry-level courses at the undergraduate program. The courses were selected based on their representation of different branches of computer science that build on different traditions, namely electronics and electrical engineering, design and art, and programming and software modeling. We refer to the courses by their overarching topics, but the sketching assignments were not necessarily representative of the whole course curriculum and were mainly included as examples of different sketching assignments traditionally found within computer science. Table 1 presents an overview of the three courses and examples of assignments.

Table 1. Overview of the courses and sketching assignments

Course topic	Extracts from the curriculum	Examples of specific models
Electronics	Sketching simple digital logic circuits	Circuit diagrams, logic gates (e.g., full adder)
Interaction design	Sketching complex models and freehand sketches	Customer journey maps, storyboards, lifecycle models
Database and software modeling	Sketching relational models and diagrams	Object-role modeling, UML diagrams (e.g., sequence diagram)

As mentioned in the introduction, we did not normalize the assignments from the three courses to avoid any important details of each course disappearing while attempting to align the complexity and degree of freedom. One such example is the freehand assignments in the interaction design course where we wanted to capture underlying issues, e.g., multi-gestures with the graphical tablet used to perform complicated tasks that have previously been discussed as a source of confusion (Rante et al., 2018, p. 5).

3.2 Performance Testing and Post-Test Interview

The performance testing was conducted at the university premises and included testing stations with separate test setups for each of the input device to allow parallel testing. See Figure 1 for examples of the testing stations. Each of the 77 participants solved assignments with the three input devices for one specific course. The test was conducted in this manner due to the participants having different backgrounds and familiarity with the different courses. Recruiting participants who were equally proficient in the three subjects would have proved challenging as the course combination is not representative of courses typically taken in combination during an undergraduate degree at this university. This decision had a positive impact on the learning effects on using the three input devices as each participant only used each device once during the test. The 77 students were recruited sequentially to reach an equal distribution between the three courses and ended with 25, 24, and 28 participants in the three courses.

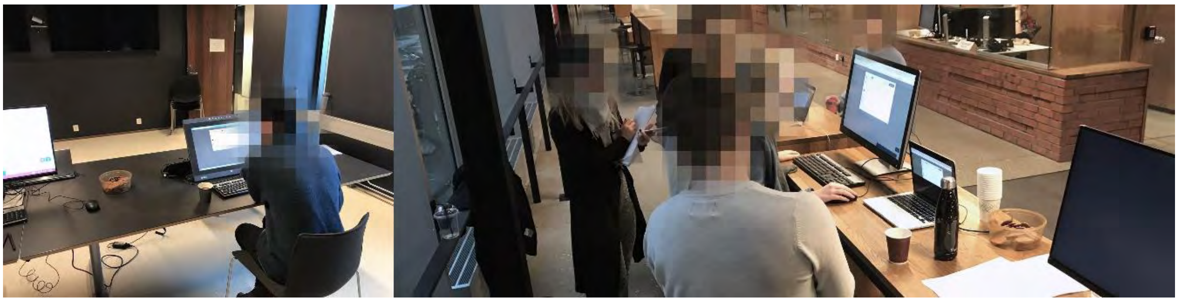

Figure 1. Setup of two testing stations

To ensure familiarity with the exam format and assignments, and to avoid bias in the method (Walker and Handley, 2016, p. 5), we utilized a nation-wide adopted online e-exam software in this study. 69 of the 77 participating students (89.6 %) had prior experience with digital exams using this specific e-exam software. All participants reported having experience with both a mouse and a touchpad while 36 of the 77 participants (46.8 %) had tried a graphical tablet before. These 36 participants were distributed evenly across the three courses.

Each participant was observed while solving the assignment, and then later questioned in a post-test interview. Screenshots were taken of each sketch. While visual analyses of the sketches were not a part of this study, it could be subject to a complementing follow-up study similar to (Taele and Hammond, 2016).

Interesting observations made during the performance tests were included in the post-test interviews to corroborate any suspicions of underlying issues.

The order of the three input devices and the assignment order were both randomized at the participant level to minimize learning bias. The participants solved different assignments within each course, and we used time-to-complete as the metric. A similar approach has been adopted by Oviatt et al. (2006, p. 5) when comparing different input devices, and Landay and Myers (2001, p. 62) also argue that task completion can be used as a measure for efficiency in such contexts.

4. RESULTS AND FINDINGS

4.1 Statistical Analysis

The performance tests resulted in three data points on performance time for each of the 77 students, separated into the three courses. We ran a two-way factorial ANOVA to determine the effects of the two independent variables *course* (electronics, interaction design, database) and *input device* (mouse, graphical tablet, touchpad) on the dependent variable completion time. All effects were statistically significant at the 0.05 level. The main effect of the course on the completion time, $F(2, 222) = 31.06$, $p < .05$, $\eta_p^2 = .219$, confirmed our initial suspicion that our sample size and research design would not support any cross-course comparisons of the three input devices due to variation in the courses, i.e., the assignments given and the complexity involved. This observation was further supported by the effect size of course ($\eta_p^2 = .219$) being larger than the effect size for the input device ($\eta_p^2 = .154$). However, our study continued with the outlined emphasis on the variation observed within each course based on the input device. The main effects of course and input device were qualified by a statistically significant interaction effect on completion time, $F(4, 222) = 4.43$, $p < .05$, $\eta_p^2 = .074$. To understand the exact nature of the interaction, we continued with a Bonferroni-adjusted test of simple effects. The comparisons indicated the course had a nonsignificant impact on performance when using the mouse as input device, $F(2, 222) = 3.56$, $p = .31$, while the course did impact the performance for both graphical tablet ($F(2, 222) = 7.70$, $p < .05$) and touchpad ($F(2, 222) = 28.66$, $p < .05$) users. Figure 2 (left) illustrates how the marginal means in performance time (seconds) within each course differed with an almost similar pattern, i.e., the touchpad yielding the highest time within each course and being statistically different from the homogenous subset of mouse and graphical tablet. The heat map in Figure 2 (right) visualizes a corresponding pattern with emphasis on the distribution of mean performance time.

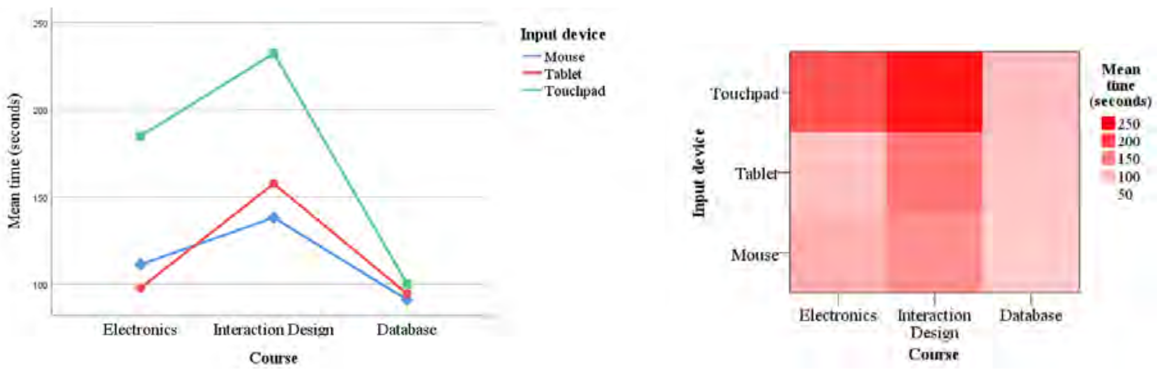

Figure 2. The mean performance time (left) and the distribution (right) for each input device across the three courses

4.2 Post-Test Interviews

While not evident from the statistical analysis alone, the qualitative data from the post-test interviews with the participants revealed further insight into the results (referred to as P1-P77). For instance, the participants explained the coinciding performance time observed as a direct result of using the mouse. The explanations were almost unanimously reported as caused by two overarching factors, namely familiarity and forced

simplicity. Familiarity in this context refers to the general experience the participants had when using a mouse for general purposes, which yielded knowledge of its interaction mechanisms and ranges of use. This experience was highly transferrable to the e-exam situation (P31, P70). Forced simplicity reflects the general knowledge people had gained using a mouse in the past performing similar tasks, which automatically balanced the complexity of the sketch according to prior experience with how a mouse performs during such tasks. As an example of forced simplicity, we observed very few instances of overcomplicating sketches for the mouse compared to the two other input devices.

An important observation made was that many students chose to enhance or comment their sketching by using built-in auxiliary elements not intended for sketching purposes, e.g., text boxes and arrows, to compensate for missing or hidden functions they were familiar with from software used in the past. Making these actions were reported as both quicker and easier in use when interacting with a mouse rather than either the graphical tablet or the touchpad. Students were also quick to make comparisons with other, more familiar alternatives and argue that specialized and adapted software would enhance the overall sketching experience. This was a recurring feedback in all courses: "*this would be best with Logisim*" (P7 on sketching of digital logic circuits in the electronics course), *"If things were like Sketch, I would have spent only half the time"* (P32 sketching a customer journey map in the interaction design course), and *"draw.io would be both faster and more intuitive"* (P53 on sketching of database and class models in the database course). This pattern was also reflected in several comments during the post-test interviews directed at the input devices' ability to interact with specific software components rather than addressing the input device itself as a sketching tool.

The details from the statistical analysis concerning each individual course suggest that the mouse and the graphical tablet performed quite equally on average. They were only matched by the touchpad in the database course where the three mean performance times were quite similar with only nonsignificant differences. The participants reported that the assignments typical for that course did not require any smaller details nor any high level of precision to convey the message behind the sketch. The general feedback for the touchpad was that it was unsuitable for most types of sketches associated with the three courses in this study. Multiple participants (P10, P12, P14, P43) went as far as to state that the touchpad was practically impossible to use for several of the assignments: *"I would rather write how it would look than handing in this sketch"* (P43). Another recurring issue most prominent in the case of the electronics and interaction design course was the interaction mechanisms of the touchpad. The limited physical size of the touchpad and the inaccuracy of the fingertip compared to a mouse prevented students from making either longer movements and gesture or precise movements needed for detailed sketches (P12, P31).

The main positive comment mentioned about the touchpad was its proximity to the keyboard and the increased efficiency for smaller, simple sketches with text elements (P29, P50). The simplistic and minimal solutions to sketching assignments in the electronic course were mainly responsible for the huge difference in performance time when comparing the electronic course versus the interaction design and database course (illustrated in Figure 3). The feedback from the students indicated that the smaller, more detailed sketching such as circuit sketches required fewer or no large movements in the physical space while simultaneously occupying a smaller portion of the screen.

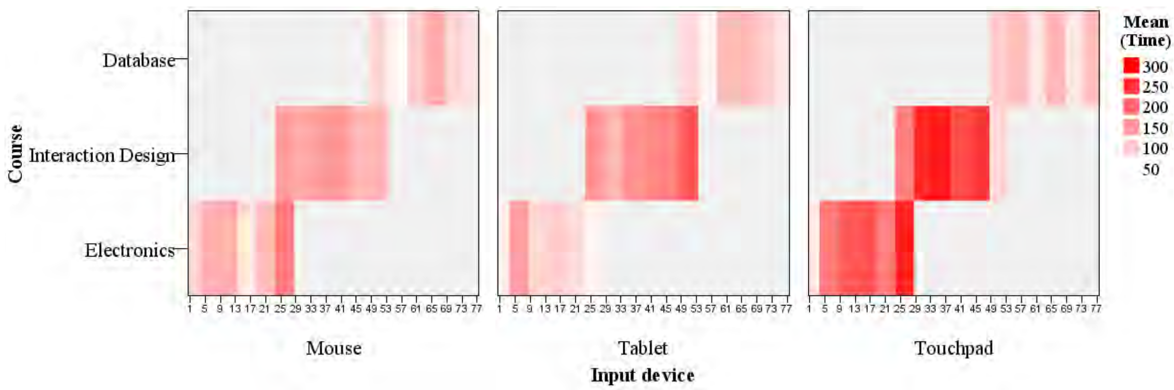

Figure 3. Distribution of mean time for each input device across the three courses

The performance using the graphical tablet yielded results similar to using the mouse in all three courses. The participants reported their experience with using the graphical tablet with mixed results. Most people preferred the graphical tablet to the two alternatives for more complex sketches, predominantly in the interaction design course, where there were less formal structures and elements involved. Regardless of the overall performance, the participants seemed pleased with the graphical tablet, and several participants even reported it as the most favorable alternative, despite not being the fastest (P8, P61, P63). Even though almost half the participants (46.8 %) had used a graphical tablet before, and a recurring feedback was that it was considered the best alternative, it still required practice within the e-exam situation with both the hardware and software available. This was especially the case for those without prior experience, as highlighted by one participant: *"I do believe that the graphical tablet would be better than a mouse, but I have not used it before. I use the mouse every day, so I have a certain level of control, but it is still a hopeless tool for sketching"* (P12).

With experience, students would be less prone to time spent on understanding how the hardware and software harmonized, and use gestures that are more advanced as part of the sketch. The lack of experience with graphical tablets is the main explanation for the spike in performance time for the interaction design course (Figure 2, left). However, experience with the graphical tablet alone would not solve all issues. The participants who had used graphical tablets in the past, especially those considering themselves proficient users, experienced dissonance between expectation and behavior. Most advanced features of a generic tablet, e.g., multi-touch gestures, pressure sensitivity, and tilt support, are not readily available in generic browser-based exam software, which limited their use also in our study. Multiple participants explained their performance with a lack of support for expected gestures and time spent trying to find interface-specific alternatives (P40, P71).

4.3 Findings

The two sets of data allowed us to study patterns of mean performance times within each course and qualify any statistically significant results with feedback received during the post-test interviews. The specific benefits and challenges of each input device themselves are not the main point of this paper, but rather how they fulfill different roles and functions in the different types of sketching assignments. As such, the absolute values of performance are of less interest than their relative value.

Both the statistical analysis and the post-test interviews suggest that one input device cannot be considered unanimously better than its alternatives, and that remains the overarching answer to our initial research question. While the performance score of the mouse was the most consistent in all three courses, the preferences of the participants suggested that the graphical tablet was the most preferred input device for several types of sketches. The touchpad rated lowest in both performance and preference except for a scope of specific types of smaller, simple sketches. However, the data also revealed specific benefits and challenges with all three input devices that can suggest opportunities and limitations relevant to other studies. Table 2 outlines the main benefits and main challenges identified for each input device.

Table 2. Overview of the main benefits and challenges for each input device

Input device	Main benefits	Main challenges
Mouse	▪ The most consistent performance for sketching assignments across the three courses ▪ Gave the best balance between expectations and attempted complexity in the sketching ▪ No significant difference in performance times for the three courses	▪ Demanded a larger physical space when sketching bigger figures that involve longer dragging movements with the mouse ▪ Preferences of the users required adjustment of sensitivity for more complex and detailed sketches

Input device	Main benefits	Main challenges
Graphical Tablet	▪ The most preferred option for larger and complex sketches ▪ Suitable for sketches involving both technical, precise elements and freehand, conceptual elements ▪ The input device perceived as most similar to classic pen-and-paper sketches	▪ Usually more time-consuming, even for proficient users with experience ▪ Less familiar as an input device in an e-exam situation for many students ▪ Heavily dependent on a software-hardware symbiosis to maximize its potential
Touchpad	▪ Can be used with students' own laptops without the need for additional devices ▪ Considered the most efficient option for smaller, simple sketches in combination with text elements	▪ Unsuitable for most situations involving both detailed, precise sketches and larger, complex sketches ▪ The small physical surface prevents long gestures which may preclude certain types of sketches ▪ Consistently the most time-consuming and frustrating input device

5. DISCUSSION

Our answer to the initial research question was that all input devices for e-exams within computer science courses that include sketching assignments involve compromises, either in terms of efficiency or in terms of the perceived experience from the students' point of view. In line with previous studies, our research design revolved around emphasizing the perspective of the student, and we have scoped our study to very specific issues rather than attempting to address the broadness of challenges associated with e-exam, as seen in, e.g., (Wibowo et al., 2016). As a result, our discussion aligns the findings with existing literature on similar studies, specifically targeting e-exams or sketching assignments.

The generic nature of the e-exam tool generated both frustrating and unfamiliar responses due to the deviation in behavior from other specialized software that the students had previously used in their studies. Building on existing practice in teaching through customized interfaces may minimize the unnecessary complexity and help the students keep their attention directed towards the actual problem-solving (Oviatt et al., 2006, p. 2). However, the strong variation in the sketch complexity and degree of freedom seen in our study would suggest that the teachers' sense of ownership and control (Motschnig et al., 2016) should be extended to include shaping or adjusting of the exam interface. A similar argument has been raised by Lazem who calls for the *"adaptation and appropriation of the methods to the local culture"* (2016, p. 8) when discussing a course with a similar curriculum as the interaction design course. However, accommodating to teachers' needs may prove challenging as teachers may require equally simplified interfaces to support their workload and work practice (Kuikka et al., 2014), especially considering teachers have reported to be evaluated in their production of learning outcomes when using online assessment against assessment based on traditional methodologies (Crisp, 2011, p. 12). This is of increased importance for sketching assignments as the teacher's ability to shape the assignments is also dependent on the physical setup of the e-exam. In our study, the physical surface area of the touchpad yielded high variation in the average mean performance time across the three courses. For certain exam setups, e.g., BYOD approaches (Walker and Handley, 2016, p. 2), using self-brought laptop with different touchpads may be of such importance to the students' performance that it may end up restricting the teacher's freedom in designing exam assignments involving sketches.

During the post-test interviews, we also discovered that several participants suggested that they suspected their own performance to be greatly increased with just minimal time to familiarize themselves with the mechanisms of the graphical tablet. The graphical tablet scored best on preference despite not being the fastest option, and participants claimed to be willing to invest time into practice if it had been a real alternative. This

is an interesting contrast to similar studies in mathematics, where the graphical tablet was disliked (Oviatt et al., 2006, p. 9). Nevertheless, the argument of motivating through and with familiarity has been previously discussed, e.g., by (Rante et al., 2018) who discuss building on learning modalities from childhood, or through the use of already familiar technologies such as mobile phones to prolong learning (Romero et al., 2009). As such, it would be interesting to follow up with additional studies where participants were allowed more training ahead of the assessment to increase their familiarity with the interface. For shorter assignments, Polsley et al. have argued that increasing time spent on problems from just one to two minutes can increase the completion rates significantly (2016, p. 109).

Another important everlasting issue of digitalized versions of traditionally analog tasks is the struggle to convey enough familiarity, safety, and security to embrace all relevant end-users. In our case, the post-test interviews revealed that 28 % of the participants still preferred pen-and-paper after testing with the three input devices. As such, our results are not as equally divided in performance and preference as for instance (Walker and Handley, 2016; Wibowo et al., 2016). While the students who participated in this study are particularly familiar with computers and arguably less prone to computer-related unease, the experience of facing unfamiliar systems, especially in the context of an exam, may still provoke or reinforce computer aversion or anxiety (Walker and Handley, 2016, p. 8). Studies have pointed to unfamiliarity and inefficiency as contributors to why some students still prefer paper-based exams over e-exams (Wibowo et al., 2016, p. 19). Our participants did not express any issues with anxiety or stress, but this was probably due to the lack of an actual exam situation. Nevertheless, the participants vividly expressed their frustration over unnecessary inefficiency. This was the case, in particular, for the touchpad. These experiences were an important contributing factor to a large number of participants that reported that they still preferred pen-and-paper during the post-test interviews. Another relevant factor that emerged in the post-test interviews was the fact that the digitalization of pen-and-paper depended on a screen, which disturbed the hand-eye coordination, a concern that several participants claimed was of utmost importance in sketching assignments. Rante et al. point to similar results where the separation of eye and hand while looking at a screen and simultaneously making moves on a tablet out of the visual range may complicate the interaction (2018, p. 4), and believe that stylus against a screen may better the results. As such, it remains to be seen whether the digitalization mimicking a pen-gesture in its common form have intrinsic challenges that may never make it a full worthy replacement for pen-and-paper in an e-exam context, at least not without becoming an entirely different experience. However, recent studies have experimented with novel interfaces, e.g., enabling augmentation of paper interfaces (Tsandilas et al., 2015) and illuminated, interactive paper (Klamka and Dachselt, 2017), and such approaches may be able to overcome the challenge experienced with all three input devices included in our study.

A final point worth mentioning concerns the challenges experienced by the participants of our study that were directly linked to the lack of coordination between the input devices and the exam tool. Issues were reported on both the software behavior and the mechanisms of the interface, but the e-exam software was not adapted to support the full range of gestures and features of the mouse, touchpad, and graphical tablet. This was notably the case for the graphical tablet as it supported the broadest range of features. Participants new to such an interface struggled with unfamiliarity due to inexperience, while proficient users of graphical tablets encountered challenges due to a mismatch between expected and actual behavior. As such, returning to the point of familiarity, we argue that a good exam tool should have a strong cohesion between how the teacher expects the students to answer assignments as well as the coordination between the hardware and software involved to support the students while doing so. Having systems for e-exams that are adaptable to different input modalities at both a hardware and software level may also overcome some of the challenges related to achieving proper inclusion and accessibility (Drake et al., 2015, p. 133).

It might be unlikely or even impossible to shape the e-exam in a manner that accommodates the preferences of all students (Walker and Handley, 2016, p. 11). However, supporting the perspective of the students through familiarity may contribute to shifting the students' perspective on e-exam tools as merely a measurement tool to something that can stimulate learning (Walters et al., 2017, p. 1163).

6. CONCLUSION

Our findings point to benefits and challenges observed and experienced with the three input devices and serve as indicators of opportunities and limitations on when and how the respective input devices could be utilized in an e-exam context. Our study only investigated sketching assignments from the field of computer science, yet the differences in both the performance and preference were significant. As such, we argue that e-exam systems intending to support sketching assignments should avoid the one generic solution and rather support a multitude of input devices. However, the input device alone is not the only source of challenges when solving sketching assignments. Issues such as familiarity, expectations, and the hardware-software harmony were among the related concerns that have been discussed to shed light on the complexity of the matter.

REFERENCES

Abdel Karim, N., Shukur, Z., 2016. Using preferences as user identification in the online examination. Int. J. Adv. Sci. Eng. Inf. Technol. 6, 1026–1032.

Barreto, L., Taele, P., Hammond, T., 2016. A Stylus-Driven Intelligent Tutoring System for Music Education Instruction, in: Hammond, T., Valentine, S., Adler, A. (Eds.), Revolutionizing Education with Digital Ink: The Impact of Pen and Touch Technology on Education, Human–Computer Interaction Series. Springer International Publishing, Cham, pp. 141–161. https://doi.org/10.1007/978-3-319-31193-7_10

Crisp, G., 2011. Teacher's Handbook on e-Assessment. Transform. Assess.- ALTC Fellowsh. Act. 18.

Do, E.Y.-L., 2002. Drawing marks, acts, and reacts: Toward a computational sketching interface for architectural design. AI EDAM 16, 149–171.

Drake, J.R., O'Hara, M., Seeman, E., 2015. Five principles for MOOC design: With a case study. J. Inf. Technol. Educ. Innov. Pract. 14, 125–143.

Ibrahim, W., Atif, Y., Shuaib, K., Sampson, D., 2015. A web-based course assessment tool with direct mapping to student outcomes. Educ. Technol. Soc. 18, 46–59.

Klamka, K., Dachselt, R., 2017. IllumiPaper: Illuminated interactive paper, in: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. ACM, pp. 5605–5618.

Koile, K., Rubin, A., 2016. Tablet-Based Technology to Support Students' Understanding of Division, in: Hammond, T., Valentine, S., Adler, A. (Eds.), Revolutionizing Education with Digital Ink: The Impact of Pen and Touch Technology on Education, Human–Computer Interaction Series. Springer International Publishing, Cham, pp. 71–89. https://doi.org/10.1007/978-3-319-31193-7_5

Kuikka, M., Kitola, M., Laakso, M.-J., 2014. Challenges when introducing electronic exam. Res. Learn. Technol. 22.

Landay, J.A., Myers, B.A., 2001. Sketching interfaces: Toward more human interface design. Computer 34, 56–64.

Lane, D., 2017. Drawing and Sketching: Understanding the Complexity of Paper-Pencil Interactions within Technology Education. Handb. Technol. Educ. 1–18.

Lara-Garduno, R., Leslie, N., Hammond, T., 2016. SmartStrokes: Digitizing Paper-Based Neuropsychological Tests, in: Hammond, T., Valentine, S., Adler, A. (Eds.), Revolutionizing Education with Digital Ink: The Impact of Pen and Touch Technology on Education, Human–Computer Interaction Series. Springer International Publishing, Cham, pp. 163–175. https://doi.org/10.1007/978-3-319-31193-7_11

Lazem, S., 2016. A case study for sensitising Egyptian engineering students to user-experience in technology design, in: Proceedings of the 7th Annual Symposium on Computing for Development. ACM, p. 12.

Lee, T.-H., Wu, F.-G., Chen, H.-T., 2017. Innovation & evaluation of tangible direct manipulation digital drawing pens for children. Appl. Ergon. 60, 207–219.

Lewis, M., Sturdee, M., Marquardt, N., 2019. Sketching in HCI: Hands-on Course of Sketching Techniques, in: Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems, CHI EA '19. ACM, New York, NY, USA, p. C04:1–C04:5. https://doi.org/10.1145/3290607.3298805

Motschnig, R., Sedlmair, M., Schröder, S., Möller, T., 2016. A team-approach to putting learner-centered principles to practice in a large course on Human-Computer Interaction, in: 2016 IEEE Frontiers in Education Conference (FIE). IEEE, pp. 1–9.

Osang, F., 2012. Electronic examination in Nigeria, academic staff perspective—Case study: National Open University of Nigeria (NOUN). Int. J. Inf. Educ. Technol. 2, 304–307.

Oviatt, S., Arthur, A., Cohen, J., 2006. Quiet interfaces that help students think, in: Proceedings of the 19th Annual ACM Symposium on User Interface Software and Technology. ACM, pp. 191–200.

Polsley, S., Ray, J., Nelligan, T., Helms, M., Linsey, J., Hammond, T., 2016. Leveraging Trends in Student Interaction to Enhance the Effectiveness of Sketch-Based Educational Software, in: Hammond, T., Valentine, S., Adler, A. (Eds.), Revolutionizing Education with Digital Ink: The Impact of Pen and Touch Technology on Education, Human–Computer Interaction Series. Springer International Publishing, Cham, pp. 103–114. https://doi.org/10.1007/978-3-319-31193-7_7

Rante, H., Lund, M., Caliz, D., 2018. The Role of Tangible Interfaces in Enhancing Children's Engagement in Learning, in: MATEC Web of Conferences. EDP Sciences, p. 01013.

Romero, C., Ventura, S., De Bra, P., 2009. Using mobile and web-based computerized tests to evaluate university students. Comput. Appl. Eng. Educ. 17, 435–447.

Taele, P., Hammond, T., 2016. An Intelligent Sketch-Based Educational Interface for Learning Complex Written East Asian Phonetic Symbols, in: Hammond, T., Valentine, S., Adler, A. (Eds.), Revolutionizing Education with Digital Ink: The Impact of Pen and Touch Technology on Education, Human–Computer Interaction Series. Springer International Publishing, Cham, pp. 129–140. https://doi.org/10.1007/978-3-319-31193-7_9

Thompson, R., Tanimoto, S., Berninger, V., Nagy, W., 2016. Design Studies for Stylus and Finger-Based Interaction in Writing Instruction on Tablets, in: Revolutionizing Education with Digital Ink. Springer, pp. 51–69.

Tsandilas, T., Grammatikou, M., Huot, S., 2015. Bricosketch: Mixing paper and computer drawing tools in professional illustration. Presented at the Proceedings of the 2015 International Conference on Interactive Tabletops & Surfaces, ACM, pp. 127–136.

Walker, R., Handley, Z., 2016. Designing for learner engagement with computer-based testing. Res. Learn. Technol. 24.

Walters, S.R., Raposo Ferreira da Silva, P.L., Nikolai, J., 2017. Teaching, learning, and assessment: insights into students' motivation to learn.

Wang, G., Bowditch, N., Zeleznik, R., Kwon, M., LaViola, J.J., 2016. A Tablet-Based Math Tutor for Beginning Algebra, in: Hammond, T., Valentine, S., Adler, A. (Eds.), Revolutionizing Education with Digital Ink: The Impact of Pen and Touch Technology on Education, Human–Computer Interaction Series. Springer International Publishing, Cham, pp. 91–102. https://doi.org/10.1007/978-3-319-31193-7_6

Wibowo, S., Grandhi, S., Chugh, R., Sawir, E., 2016. A pilot study of an electronic exam system at an Australian University. J. Educ. Technol. Syst. 45, 5–33.

Williford, B., Taele, P., Nelligan, T., Li, W., Linsey, J., Hammond, T., 2016. Persketchtivity: an intelligent pen-based educational application for design sketching instruction, in: Revolutionizing Education with Digital Ink. Springer, pp. 115–127.

Yang, M.C., Cham, J.G., 2006. An Analysis of Sketching Skill and Its Role in Early Stage Engineering Design. J. Mech. Des. 129, 476–482. https://doi.org/10.1115/1.2712214